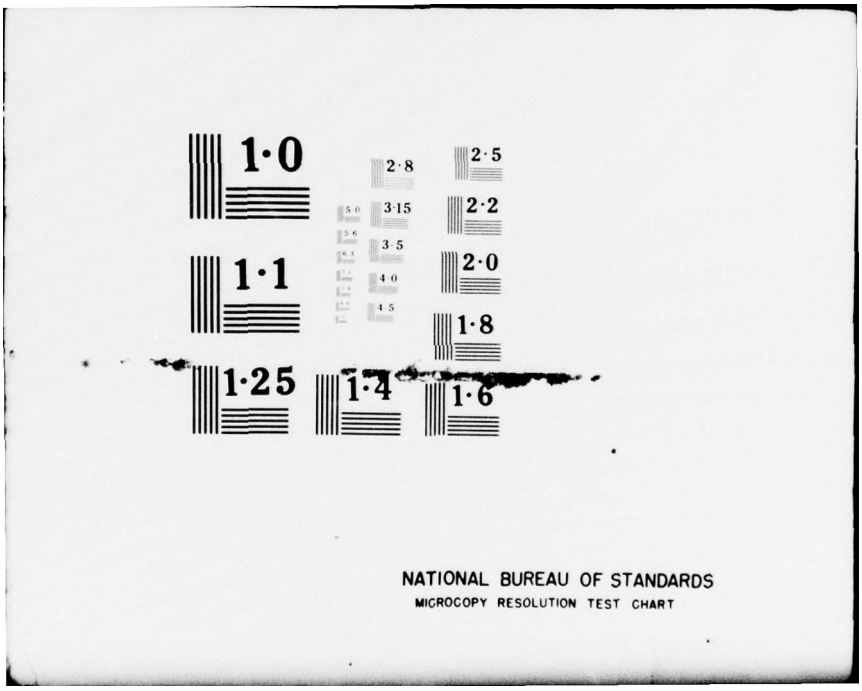
AD-A046 647 COLORADO STATE UNIV FORT COLLINS DEPT OF MATHEMATICS F/G 12/1
UNIFORM APPROXIMATION WITH RATIONAL FUNCTIONS HAVING NEGATIVE P--ETC(U)
JUN 77 E H KAUFMAN, G D TAYLOR AFOSR-76-2878

UNCLASSIFIED AFOSR-TR-77-1260 NL

1 OF 1
ADA
046647

END
DATE
FILMED
12-77
DDC


1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5
2·8
3·15
3·5
4·0
4·5
5·0
5·6
6·3
NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (19) (18) AFOSR-TR-77-1260	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) UNIFORM APPROXIMATION WITH RATIONAL FUNCTIONS HAVING NEGATIVE POLES,	5. TYPE OF REPORT & PERIOD COVERED: Interim (2) / 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) E. H. Kaufman, Jr. and G. D. Taylor (15)	8. CONTRACT OR GRANT NUMBER(s): AFOSR-76-2878
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Colorado State University, Department of Mathematics, Fort Collins, CO 80523	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F (16) 2304/A3 (17) A3
11. CONTROLLING OFFICE NAME AND ADDRESS: AFOSR/NM, Bldg. 410, Bolling AFB, D.C. 20332	12. REPORT DATE: (11) June 1977 / 13. NUMBER OF PAGES: ~~21~~ (12) 22 p.
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED / 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Uniform rational approximation, constrainted rational approximation, constrained nonlinear approximation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A general theory of uniform approximation with rational functions having negative poles is developed. An existence theory is given and local characterization and uniqueness are developed. Algorithms for computing these approximants are given, together with numerical results.

407344

AFOSR-TR- 77-1260

UNIFORM APPROXIMATION WITH RATIONAL FUNCTIONS HAVING NEGATIVE POLES

by

E. H. Kaufman, Jr. and G. D. Taylor[1]

ABSTRACT

A general theory of uniform approximation with rational functions having negative poles is developed. An existence theory is given and local characterization and uniqueness are developed. Algorithms for computing these approximants are given, together with numerical results.

1. Introduction

Let Π_m denote the space of all real algebraic polynomials of degree less than or equal to m. For m = 1, 2, ..., define $\mathcal{R}_m$ by

$$\mathcal{R}_m = \{R = P/Q : P \in \Pi_{m-1},\ Q(x) = \prod_{i=1}^{m} (q_i x + 1),\ q_i \geq 0 \text{ for all } i\}$$

and $\tilde{\mathcal{R}}_m$ by

$$\tilde{\mathcal{R}}_m = \{R = P/Q : P \in \Pi_{m-1},\ Q(x) = (qx + 1)^m,\ q \geq 0\}.$$

Thus $\mathcal{R}_m$ is the collection of all rational functions with denominator 1 or negative poles from $\mathcal{R}_m^{m-1}[0, \infty)$ and $\tilde{\mathcal{R}}_m$ is the collection of all rational functions with an m^{th} order negative pole or denominator 1 from $\mathcal{R}_m^{m-1}[0, \infty)$. Let $[0, \alpha]$ be an interval where $\alpha = \infty$ is permissible (so that $[0, \alpha] = [0, \infty)$). Let $f \in C[0, \alpha]$ where we shall assume that $\lim_{x \to \infty} f(x) = 0$ if $\alpha = \infty$. In this setting, we wish to study the following approximation theory problems: Find

$$\lambda_m(f) = \inf \{ \|f - R\|_{L^\infty[0,\alpha]} : R \in \mathcal{R}_m \} \tag{1.1}$$

and

$$\mu_m(f) = \inf \{ \|f - R\|_{L^\infty[0,\alpha]} : R \in \tilde{\mathcal{R}}_m \}. \tag{1.2}$$

[1] Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. 76-2870.

The motivation for this study is a recent paper of Saff, Schönhage and Varga [7] where it is shown that there exists a sequence $\{R_m\}_{m=1}^{\infty}$, with $R_m(x) = P_{m-1}(x)/(1 + \frac{x}{m})^m$, $P_{m-1} \in \Pi_{m-1}$ such that

$$3 - 2\sqrt{2} \leq \overline{\lim_{m\to\infty}} \| e^{-x} - R_m \|^{1/m}_{L^\infty[0,\infty)} \leq \frac{1}{2}.$$

That is, $\{R_m(x)\}$ converges geometrically to e^{-x} on $[0, \infty)$. In addition, since the poles of $R_m(x)$ are all real it follows that $R_m(z)$ must converge geometrically to e^{-z} in an infinite sector symmetric about the positive x-axis [8]. An application of this theory is in the construction of numerical solutions for solving linear systems of ordinary differential equations which arise from semi-discretization of linear parabolic partial differential equations (see [1] and [7]). For example, as described in [7], consider the numerical solution of the linear system of ordinary differential equations

$$\frac{d\underset{\sim}{u}(t)}{dt} = -A\underset{\sim}{u}(t) + \underset{\sim}{k}, \quad t > 0 \tag{1.3}$$

$$\underset{\sim}{u}(0) = \underset{\sim}{u}_0$$

where $\underset{\sim}{u}(t) = [u_1(t), \ldots, u_n(t)]^T$ is a column vector with n components and A is an $(n\times n)$ positive definite symmetric matrix. The integer n is related to the mesh size of the discretization and can be large. The solution to (1.3) is given explicitly by

$$\underset{\sim}{u}(t) = A^{-1}\underset{\sim}{k} + \exp(-tA)\{\underset{\sim}{u}_0 - A^{-1}\underset{\sim}{k}\} \tag{1.4}$$

for all $t \geq 0$, where $\exp(-tA) \doteq \sum_{\nu=0}^{\infty} (-tA)^\nu/\nu!$. For computational purposes one must approximate $\exp(-\Delta tA)$. In [7] this is done by using $R_m(\Delta tA) \doteq (I + \frac{\Delta t}{m}A)^{-m} P_{m-1}(\Delta tA)$ where $P_{m-1} \in \Pi_{m-1}$ is the solution to $\inf\{ \| e^{-x} - \frac{P(x)}{(1 + \frac{x}{m})^m} \|_{L^\infty[0,\infty)} : P \in \Pi_{m-1}\} = \rho_m$. That such a P_{m-1} exists

and is unique follows from the theory of best uniform approximation with Haar subspaces; it can be calculated via the standard Remes algorithm if one works on [0, b] with b sufficiently large. One then computes approximations $\underset{\sim}{w}^{(r)}$ to $\underset{\sim}{u}(r\Delta t)$ for r = 1, 2, ..., where $\underset{\sim}{w}^{(0)} = \underset{\sim}{u}_0$ and

$$\underset{\sim}{w}^{(r)} \doteq A^{-1}\underset{\sim}{k} + R_m(\Delta tA)\{\underset{\sim}{w}^{(r-1)} - A^{-1}\underset{\sim}{k}\}. \tag{1.5}$$

Due to the special form of the denominator of R_m, $\underset{\sim}{w}^{(r)}$ can be obtained from the repeated inversion of

$$(I + \frac{\Delta t}{m}A)\underset{\sim}{g}_{\ell+1} = \underset{\sim}{g}_\ell, \quad 0 \le \ell \le m - 1 \tag{1.6}$$

m times with $\underset{\sim}{g}_0 \doteq (I + \frac{\Delta t}{m}A)^m A^{-1}\underset{\sim}{k} + P_{m-1}(\Delta tA)\{w^{(r-1)} - A^{-1}k\}$. Numerically, this method is attractive in that an LU factorization can be done for $I + \frac{\Delta t}{m}A$ once and then $\underset{\sim}{g}_m = \underset{\sim}{w}^{(r)}$ can be calculated by performing a forward substitution followed by a backward substitution m times. In addition, the matrix $I + \frac{\Delta t}{m}A$ will be a band matrix since A will have a band structure inherited from the finite difference formulas used.

Thus, one is motivated to construct a similar numerical method built around a "solution" $R_m^* \in \mathcal{R}_m$ to (1.1). Hopefully, the increased accuracy of approximating e^{-x} with R_m^* will allow for a smaller choice of m in $\frac{\Delta t}{m}$. The apparent disadvantage of this method compared to that described above is that $\underset{\sim}{w}^{(r)}$ is now found by solving

$$\{\prod_{i=1}^{m}(I + q_i\Delta tA)\}\underset{\sim}{w}^{(r)} = \{\prod_{i=1}^{m}(I + q_i\Delta tA)\}A^{-1}\underset{\sim}{k} + P_{m-1}^*(\Delta tA)\{\underset{\sim}{w}^{(r-1)} - A^{-1}\underset{\sim}{k}\} \tag{1.7}$$

which will involve increased computation, where $R_m^*(x) = P_{m-1}^*(x) / \prod_{i=1}^{m} (q_i x + 1)$. We say apparent disadvantage since our numerical results suggest that $R_m^* \in \tilde{\mathcal{R}}_m$. That is, R_m^* appears to give a rise to the same sort of method as corresponding to R_m with increased accuracy for no additional effort. In fact, this is known to be true in theory also, for the case that m=2 [4] (R. S. Varga has informed us that this has also been done independently by A. Schönhage). We will return to this case later.

In the next two sections we shall prove an existence theorem and local characterization and uniqueness results for both R_m and $\tilde{\mathcal{R}}_m$, and consider the special case of approximating e^{-x} on $[0, \infty)$ from $\tilde{R}_m$. Then we shall describe an algorithm for computation with these spaces and give some numerical results. Finally, we will close the paper with a listing of some open problems.

2. Theoretical Results

We begin this section with a proof of existence of best approximations from $\mathcal{R}_m$ for each $f \in C[0, \alpha]$.

Theorem 2.1. Fix $f \in C[0, \alpha]$ then there exists $R^* \in \mathcal{R}_m$ for which $\|f - R^*\| = \inf\{\|f - R\| : R \in \mathcal{R}_m\}$ where $\|\cdot\| \equiv \|\cdot\|_{L^\infty[0,\alpha]}$.

Proof: Let us assume that $f \notin \mathcal{R}_m$, $0 \in \mathcal{R}_m$ is not a best approximation of f (i.e., $\|f\| > \lambda_m(f)$) and that $\max\{f(x) : x \in [0, \alpha]\} = \|f\|$. Note, if this last condition is not met then we replace f by -f and proceed as below. Thus, there exists a closed interval [a, b], b > a, such that $[a, b] \subset (0, \alpha]$, b is finite and $\min\{f(x) : x \in [a, b]\} = \gamma > \lambda_m(f)$. Let $x_0 = \frac{a+b}{2}$ and select $\{R_k\}_{k=1}^{\infty} = \{P_k/Q_k\}_{k=1}^{\infty} \subset \mathcal{R}_m$ such that

$$\frac{\gamma + \lambda_m}{2} \geq \|f - R_k\| \to \lambda_m \text{ as } k \to \infty \quad (\lambda_m = \lambda_m(f)).$$ Thus,

$$(2.1) \qquad f(x) - \frac{\gamma + \lambda_m}{2} \leq R_k(x) \leq f(x) + \frac{\gamma + \lambda_m}{2}$$

for all $x \in [0, \alpha]$. Now, let us normalize $R_k(x)$ by requiring that $Q_k(x) = \prod_{i=1}^{m} (q_i^{(k)}(x - x_0) + 1)$ where $0 \leq q_i^{(k)} < \frac{1}{x_0}$. Note this can be done since $Q_k(x)$ is known to have all negative roots. Thus, if $Q_k(x) = \prod_{i=1}^{p} (x - r_i)$, $p \leq m$ and $r_i < 0$ for all $i = 1, \ldots, p$ then we may rewrite it as

$$Q_k(x) = \left(\prod_{i=1}^{p} (x_0 - r_i) \right) \prod_{i=1}^{p} \left[\left(\frac{1}{x_0 - r_i} \right)(x - x_0) + 1 \right].$$

Set $q_i^{(k)} = \frac{1}{x_0 - r_i}$ for $i = 1, \ldots, p$, $q_i^{(k)} = 0$ for $i = p + 1, \ldots, m$ and note that $r_i < 0$ implies that $0 < q_i^{(k)} < \frac{1}{x_0}$ for $i = 1, \ldots, p$. Finally, the constant $\left(\prod_{i=1}^{p} (x_0 - r_i) \right)^{-1}$ is to be incorporated into $P_k(x)$.

Since f is bounded, we have from (2.1) that there exists a constant $M > 0$ independent of k such that for all $x \in [0, \alpha]$

$$(2.2) \qquad -M \leq R_k(x) \leq M.$$

Since $\{q_i^{(k)}\}_{i=1,k=1}^{m,\infty} \subset [0, \frac{1}{x_0})$, we may extract convergent subsequences (relabelling) such that $q_i^{(k)} \to q_i \in [0, \frac{1}{x_0}]$ for $i = 1, \ldots, m$. Note that if $q_i = \frac{1}{x_0}$ then $q_i(x - x_0) + 1$ reduces to $\frac{x}{x_0}$. Thus, $Q_k(x)$ converges uniformly to Q on compact subsets of the real line. Now, (2.1) restricted to [a, b] gives that there exist constants c_1, c_2 both positive and independent of k such that

$$(2.3) \qquad c_1 Q_k(x) \leq P_k(x) \leq c_2 Q_k(x)$$

for all $x \in [a,b]$. Now, for $x \in [a,b]$ and $q_i^{(k)} \in [0, \frac{1}{x_0}]$ we have that

$$q_i^{(k)}(x - x_0) + 1 \leq \frac{1}{x_0}(b - \frac{a + b}{2}) + 1 = \frac{2b}{b + a}$$

and

$$q_i^{(k)}(x - x_0) + 1 \geq \frac{1}{x_0}(a - \frac{a + b}{2}) + 1 = \frac{2a}{b + a} .$$

Thus, by (2.3) there exist positive constants A_1, A_2 independent of k ($A_1 = c_1(\frac{2a}{b + a})^m$, $A_2 = c_2(\frac{2b}{b + a})^m$) such that

$$A_1 \leq P_k(x) \leq A_2 \tag{2.4}$$

for all $x \in [a, b]$. By compactness of bounded subsets of Π_{m-1}, we may by (2.4) extract a convergent subsequence of $\{P_k\}$ (relabelling) such that $P_k \to P \in \Pi_{m-1}$ uniformly on compact subsets of the real line. By (2.2), we have for all $x \in (0, \alpha]$, that

$$-M \leq \frac{P(x)}{Q(x)} \leq M \tag{2.5}$$

since $Q(x) > 0$ for each $x \in (0, \alpha]$. But (2.5) implies that P(x) must have the same (or greater) order of root at 0 as Q(x). Thus, Q(x) can have at most m - 1 factors of $\frac{x}{x_0}$ and upon cancelling out common factors of P and Q, we have that the resultant $P/Q \in \mathcal{R}_m$. Also, for $x \in (0, \alpha]$,

$$|f(x) - \frac{P(x)}{Q(x)}| = \lim_{k\to\infty}|f(x) - \frac{P_k(x)}{Q_k(x)}| \leq \lim_{k\to\infty} \|f - R_k\| = \lambda_m.$$

Thus, by continuity $\|f - R\| \leq \lambda_m$, $R = P/Q \in \mathcal{R}_m$ completing the proof. ∎

The same proof also establishes

Corollary 2.2. Let $f \in C[0, \alpha]$ then there exists $R^* \in \mathcal{R}_m$ such that $\|f - R^*\| = \mu_m(f)$.

Also, we would like to observe that existence when $\alpha = \infty$ actually holds for all $f \in C[0, \infty)$ for which $\lim_{x\to\infty} f(x)$ exists and is finite by the

above proof. This is so since in the case $\lim_{x\to\infty} f(x) = \|f\|$ and $|f(x)| < \|f\|$ for all $x \in [0, \infty)$ then for n sufficiently large the interval $[n, n+1]$ can be used for the interval $[a, b]$ provided 0 is not a best approximation to f.

Now, we wish to study the space $\tilde{\mathcal{R}}_m$. In what follows we shall prove a local characterization and local uniqueness theorem for this space.

Definition 2.3. $R(x) = (p_1 + p_2x + \ldots + p_mx^{m-1})/(qx+1)^m \in \tilde{\mathcal{R}}_m$ is a local best approximation to $f \in C[0, \alpha]$ on $[0, \alpha]$ if there exists $\delta > 0$ such that if $\bar{R}(x) = (\bar{p}_1 + \bar{p}_2x + \ldots + \bar{p}_mx^{m-1})/(\bar{q}x+1)^m \in \tilde{\mathcal{R}}_m$ and $|\bar{p}_i - p_i| < \delta$, $i = 1, \ldots, m$ and $|\bar{q} - q| < \delta$ then $\|\bar{f}(x)-R(x)\| \leq \|\bar{f}(x)-\bar{R}(x)\|$. In addition, if strict inequality holds whenever $\bar{R}(x) \neq R(x)$ then R is said to be locally unique.

Before we can prove our local characterization theorem, we must prove two lemmas. The first lemma states that $\tilde{\mathcal{R}}_m$ has a local Hermite solvency property of order one.

Lemma 2.4. Suppose $\bar{R}(x) = \bar{P}(x)/\bar{Q}(x) = (\bar{p}_1 + \bar{p}_2x + \ldots + \bar{p}_mx^{m-1})/(\bar{q}x+1)^m \in \tilde{\mathcal{R}}_m$ is nondegenerate (i.e. $\bar{P} \not\equiv 0$ and $\bar{P}$ and $\bar{Q}$ have no common factors) and $\bar{q} > 0$. Let $0 \leq m_1 \leq m_2$ and $m_1 + m_2 = m + 1$. Suppose $\{\bar{y}_i\}_{i=1}^{m_2} \subset [0, \alpha]$ with $\bar{y}_i < \bar{y}_{i+1}$ for all i and $\{i_1, \ldots, i_{m_1}\} \subset \{1, \ldots, m_2\}$. Then there exist $\delta > 0$ and $\delta_1 > 0$ such that if $|y_i - \bar{y}_i| \leq \delta_1$ and $|z_i - \bar{R}(\bar{y}_i)| \leq \delta_1$ for $i = 1, \ldots, m_2$, $|y'_{i_j} - \bar{y}_{i_j}| \leq \delta_1$ and $|z'_j - \bar{R}'(\bar{y}_{i_j})| \leq \delta_1$ for $j = 1, \ldots, m_1$ then there exists exactly one $R(x) = P(x)/Q(x) = (p_1+p_2x+\ldots p_mx^{m-1})/(qx+1)^m \in \tilde{\mathcal{R}}_m$ with $|p_\nu - \bar{p}_\nu| \leq \delta$, $\nu = 1, \ldots, m$, $|q - \bar{q}| \leq \delta$, $R(y_i) = z_i$, $i = 1, \ldots, m_2$ and $R'(y'_{i_j}) = z'_j$, $j = 1, \ldots, m_1$. Furthermore, with the above restrictions $p_1, \ldots, p_m, q$ depend continuously on the variables y_i, y'_{i_j}, z_i, z'_j.

Proof: This result follows from an application of the Implicit Function Theorem. Thus, one forms the system $f_\mu(\underset{\sim}{a}) = 0$, $\mu = 1, \ldots, m+1$ where $\underset{\sim}{a} = (p_1, \ldots, p_m, q, y_1, \ldots, y_{m_2}, y'_1, \ldots, y'_{m_1}, z_1, \ldots, z_{m_2}, z'_1, \ldots, z'_{m_1})$, $f_\mu(\underset{\sim}{a}) = p_1 + \ldots + p_m y_\mu^{m-1} - (qy_\mu + 1)^m z_\mu$, $\mu = 1, \ldots, m_2$ and $f_{m_2+\mu}(\underset{\sim}{a}) = (qy'_\mu + 1)(p_2 + \ldots + (m-1)p_m(y'_\mu)^{m-2}) - mq(p_1 + \ldots p_m(y'_\mu)^{m-1}) - (qy'_\mu + 1)^{m+1} z'_\mu$, $\mu = 1, \ldots, m_1$. Observe that the point $\underset{\sim}{a}_0 = (\bar{p}_1, \ldots, \bar{p}_m, \bar{q}, \bar{y}_1, \ldots, \bar{y}_{m_2}, \bar{y}_{i_1}, \ldots, \bar{y}_{i_{m_1}}, \bar{R}(\bar{y}_1), \ldots, \bar{R}(\bar{y}_{m_2}), \bar{R}'(\bar{y}_{i_1}), \ldots, \bar{R}'(\bar{y}_{i_{m_1}}))$ satisfies this system. Thus, since each function of this system has continuous first partials with respect to each of the variables of $\underset{\sim}{a}$ (or components) we need only prove that Jacobian, $J(\underset{\sim}{a})$, of the system with respect to $p_1, \ldots, p_m, q$ has a nonzero determinant at $\underset{\sim}{a} = \underset{\sim}{a}_0$. Now, by using the equalities $z_i = \bar{R}(\bar{y}_i)$ and $z'_j = \bar{R}'(\bar{y}_{i_j})$ and adding $(m+1)\bar{q}$ times the ν^{th} row to the μ^{th} row where $\mu > m_2$ and $\bar{y}_\nu = \bar{y}_{i_{\mu-m_2}}$, $\det(J(\underset{\sim}{a}_0))$ becomes

$$\begin{vmatrix}
1 & \bar{y}_1 & \cdots & \bar{y}_1^{m-1} & -m\bar{y}_1 \dfrac{\bar{P}(\bar{y}_1)}{(\bar{q}\bar{y}_1+1)} \\
\vdots & \vdots & & \vdots & \vdots \\
1 & \bar{y}_{m_2} & \cdots & \bar{y}_{m_2}^{m-1} & -m\bar{y}_{m_2} \dfrac{\bar{P}(\bar{y}_{m_2})}{(\bar{q}\bar{y}_{m_2}+1)} \\
\bar{q} & \bar{q}\bar{y}_{i_1}+(\bar{q}\bar{y}_{i_1}+1) & \cdots & \bar{q}\bar{y}_{i_1}^{m-1}+(m-1)\bar{y}_{i_1}^{m-2}(\bar{q}\bar{y}_{i_1}+1) & -m\bar{y}_{i_1}\bar{P}'(\bar{y}_{i_1})-m\bar{P}(\bar{y}_{i_1}) \\
\vdots & \vdots & & \vdots & \vdots \\
\bar{q} & \bar{q}\bar{y}_{i_{m_1}}+(\bar{q}\bar{y}_{i_{m_1}}+1) & \cdots & \bar{q}\bar{y}_{i_{m_1}}+(m-1)\bar{y}_{i_{m_1}}^{m-2}(\bar{q}\bar{y}_{i_{m_1}}+1) & -m\bar{y}_{i_{m_1}}\bar{P}'(\bar{y}_{i_{m_1}})-m\bar{P}(\bar{y}_{i_{m_1}})
\end{vmatrix}$$

Assuming that $m_1 > 0$, replace $\bar{y}_{i_1}$ in the m_2+1^{st} row by t and set G(t) equal to the resulting function of t. Note that $G \in \Pi_{m-1}$ and $G(\bar{y}_{i_1})$ is $\det(J(\underset{\sim}{a}_0))$. Define H(t) by H(t) is $\det(J(\underset{\sim}{a}_0))$ with the m_2+1^{st} row replaced by

$((\bar{q}t + 1), t(\bar{q}t + 1), \ldots, t^{m-1}(\bar{q}t + 1), -mt\bar{P}(t))$. Note that $H \in \Pi_m$, $H'(t) = G(t)$ and $H(\bar{y}_\nu) = 0$, $\nu = 1, \ldots, m_2$, $H'(\bar{y}_{i_j}) = 0$, $j = 2, \ldots, m_1$ so that H has m zeros counting multiplicities. Thus, if $H \not\equiv 0$ then H can have no more zeros. Hence, if we can show $H \not\equiv 0$ then it will follow that $H'(\bar{y}_{i_1}) \neq 0$ and so $\det(J(\underset{\sim}{a}_0)) \neq 0$ as desired. Now $H(-\frac{1}{\bar{q}}) = (-1)^{m+m_2} \frac{m}{\bar{q}}\bar{P}(-\frac{1}{\bar{q}})D$ where D is the determinant obtained from $\det(J(\underset{\sim}{a}_0))$ by deleting the $m+1^{st}$ column and the m_2+1^{st} row. Since $\bar{P}$ and $\bar{Q}$ have no common factors we have that $\bar{P}(-\frac{1}{\bar{q}}) \neq 0$. But now adding $(-\bar{q})$ times the ν^{th} row to the μ^{th} row where $\mu > m_2$ and $\bar{y}_\nu = \bar{y}_{i_{\mu-m_2}}$ (note the row containing $\bar{y}_{i_1}$ is gone from D, this is applied to the rows containing $\bar{y}_{i_2}, \ldots, \bar{y}_{i_{m_1}}$) and then factoring out $(\bar{q}\bar{y}_{i_\mu} + 1)$ from row $m_2 + \mu - 1$, $\mu = 2, \ldots, m_1$ shows that D equals a nonzero constant times a determinant which is known to have a nonzero value. A similar proof works for the case that $m_1 = 0$ (no derivatives present). In this case one simply replaces $\bar{y}_1$ in the first row by t and proceeds as above without referring to derivatives. Finally, to guarantee that $R \in \tilde{\mathcal{C}}_m$ we require that $\delta < |\bar{q}|$. ∎

Lemma 2.4 gives a pointwise local solvency property when $m_1 = 0$, pointwise in the sense that the δ_1 and δ depend upon the points at which the functions are being evaluated. In order to prove the necessity of our local characterization we need the following zero-counting property. Here we shall assume that $\hat{\alpha} > 0$ is finite, $\hat{\alpha} \leq \alpha$.

Lemma 2.5. Let $R(x) = \bar{P}(x)/\bar{Q}(x) = (\bar{p}_1 + \ldots + \bar{p}_m x^{m-1})/(\bar{q}x + 1)^m \in \tilde{\mathcal{C}}_m$ be nondegenerate and $\bar{q} > 0$. Suppose $0 \leq \bar{y}_1 < \bar{y}_2 < \ldots < \bar{y}_m \leq \hat{\alpha}$ and $\bar{y} \in [0, \alpha] \sim \{\bar{y}_1, \ldots, \bar{y}_m\}$. Let $\delta_1 > 0$ and $\delta > 0$ be chosen correspondin

to $\bar{P}/\bar{Q}$ and the point set $\{\bar{y}_1, \ldots, \bar{y}_m, \bar{y}\}$ according to Lemma 2.4 with $m_1 = 0$. Finally, for all σ, $|\sigma| \leq \delta_1$, let $R_\sigma(x) = P_\sigma(x)/Q_\sigma(x) = (p_{1\sigma} + \ldots + p_{m\sigma}x^{m-1})/(q_\sigma x + 1)^m$ be the unique function $\in \tilde{\mathcal{R}}_m$ which satisfies $R_\sigma(\bar{y}_i) = \bar{R}(\bar{y}_i)$, $i = 1, \ldots, m$, $R_\sigma(\bar{y}) = \bar{R}(\bar{y}) + \sigma$, $|p_{i\sigma} - \bar{p}_i| \leq \delta$, $i = 1, \ldots, m$ and $|q_\sigma - q| \leq \delta$. Then there exists $\delta_2 > 0$ such that if $0 < |\sigma| \leq \delta_2$, then the only zeros of $R_\sigma - \bar{R}$ in $[0, \hat{\alpha}]$ are $\bar{y}_1, \ldots, \bar{y}_m$ and $R_\sigma - \bar{R}$ changes sign at each of these that are in $(0, \hat{\alpha})$.

Proof: Suppose not. Then there exists $\sigma_j \to 0$, $\sigma_j \neq 0$ for all j such that $R_{\sigma_j} - \bar{R}$ either has an additional zero at $y_{\sigma_j} \in [0, \hat{\alpha}] \sim \{\bar{y}_1, \ldots, \bar{y}_m\}$ or $R_{\sigma_j} - \bar{R}$ fails to change sign at one of the points $\bar{y}_\nu \in (0, \hat{\alpha})$. In both cases we write y_{σ_j} for the additional zero with the understanding that $y_{\sigma_j} = \bar{y}_\ell$ for some ℓ means that $R_{\sigma_j} - \bar{R}$ does not change sign at $\bar{y}_\ell$ in this case. By passing to a subsequence, we may assume that $y_{\sigma_j} \to y^* \in [0, \hat{\alpha}]$ where here we are using our assumption that $\hat{\alpha}$ is finite. We now consider two cases.

CASE 1. $y^* \notin \{\bar{y}_i\}_{i=1}^m$. Choose $\delta_1^* > 0$ and $\delta^* > 0$ corresponding to $\bar{P}/\bar{Q}$ and the point set $\{\bar{y}_1, \ldots, \bar{y}_m, y^*\}$ according to Lemma 1 with $m_1 = 0$. Then for j sufficiently large we have that $|y_{\sigma_j} - y^*| \leq \delta_1^*$, $|R_{\sigma_j}(y_{\sigma_j}) - \bar{R}(y^*)| \leq \delta_1^*$, $|p_{i\sigma_j} - \bar{p}_i| \leq \delta^*$, for $i = 1, \ldots, m$ and $|q_{\sigma_j} - \bar{q}| \leq \delta^*$ since the parameters $p_{1\sigma}, \ldots, p_{m\sigma}, q$ depend continuously upon the remaining parameters (so that R_{σ_j} converges uniformly to $\bar{R}$ on compact subsets of $[0, \alpha)$). Now, R_{σ_j} and $\bar{R}$ agree at the points $\bar{y}_1, \ldots, \bar{y}_m, y_{\sigma_j}$ so that by the uniqueness part of Lemma 2.4 we must have $R_{\sigma_j} \equiv \bar{R}$ which is a contradiction since $R_{\sigma_j}(\bar{y}) \neq \bar{R}(\bar{y})$.

CASE 2. Suppose $y^* = \bar{y}_\ell$ for some ℓ. Choose $\delta_1^* > 0$ and $\delta^* > 0$ corresponding to the point set $\{\bar{y}_1, \ldots, \bar{y}_m\}$ according to Lemma 2.4 with $m_1 = 1$ and $i_1 = \ell$.

Then for j sufficiently large we have that $|y_{\sigma_j} - \bar{y}_\ell| \le \delta_1^*$, $|\bar{R}'(y)-\bar{R}'(\bar{y}_\ell)| \le \delta_1^*$ for all y in the closed interval I_j with endpoints y_{σ_j} and $\bar{y}_\ell$, $|p_{i\sigma_j} - \bar{p}_i| \le \delta^*$, $i = 1, \ldots, m$ and $|q_{\sigma_j} - \bar{q}| \le \delta^*$. Now from the fact that $R'_{\sigma_j} - \bar{R}'$ is continuous on I_j and $R_{\sigma_j} - \bar{R}$ vanishes at y_{σ_j} and $\bar{y}_\ell$ we have by Rolle's Theorem that $R'_{\sigma_j} - \bar{R}'$ vanishes at some point $y'_{\ell j} \in I_j$ provided $y_{\sigma_j} \neq \bar{y}_\ell$. If $y_{\sigma_j} = \bar{y}_\ell$ (for some j) then $\bar{y}_\ell \in (0, \hat{\alpha})$ and we have that $R'_{\sigma_j} - \bar{R}'$ is zero at $\bar{y}_\ell$ since $R_{\sigma_j} - \bar{R}$ does not change sign at this point in this case. Thus, $y'_{\ell j} \in I_j$ with $R'_{\sigma_j}(y'_{\ell j}) = \bar{R}'(y'_{\ell j})$ in either case. Also, $|y'_{\ell j} - \bar{y}_\ell| \le \delta_1^*$ and $|R'_{\sigma_j}(y'_{\ell j}) - \bar{R}'(\bar{y}_\ell)| = |\bar{R}'(y'_{\ell j}) - \bar{R}'(\bar{y}_\ell)| \le \delta_1^*$. Thus, by the uniqueness part of Lemma 2.4 we must have that $R_{\sigma_j} \equiv \bar{R}$ for j sufficiently large since these functions agree at $\bar{y}_1, \ldots, \bar{y}_m$ and their derivatives agree at $y'_{\ell j}$, which is our desired final contradiction. ∎

With these results we are now ready to prove our local characterizing theorem which is an alternation-type result.

Theorem 2.6. Let $m > 0$. Then a nondegenerate $\bar{R}(x)=(\bar{p}_1+\ldots+\bar{p}_m x^{m-1})/(\bar{q}x+1)^m \in \tilde{\mathcal{R}}_m$ with $\bar{q} > 0$ is a local best approximation to $f \in C[0, \alpha]$ on $[0, \alpha]$ from $\tilde{\mathcal{R}}_m$ if and only if the error curve $E(x) = f(x) - \bar{R}(x)$ has at least $m + 2$ alternating extreme points. (If $\alpha = \infty$, then we require $\lim_{x\to\infty} f(x) = 0$).

Proof: The necessity of this alternation now follows by the arguments of Theorem 7.3 [6, pp. 10-12] for varisolvent functions ($m + 1$ is the number corresponding to the degree of varisolvence there). Lemma 2.5, above, is needed for constructing a better approximation than $\bar{R}$ when $\bar{R}$ has less than $m + 2$ alternating extreme points. If 0 and α are both extreme points, a straightforward extension of Lemma 2.5 may be needed. For the case that $\alpha = \infty$, we note that since both f and $\bar{R}$ tend to 0 as $x\to\infty$ we may replace $[0,\infty)$ by $[0,\hat{\alpha}]$, $\hat{\alpha}$ finite such that for $x \ge \hat{\alpha}$, $|f(x)| + |\bar{R}(x)| \le \frac{\mu_m(f)}{4}$

(assuming $\mu_m(f) > 0$, i.e., $f \notin \tilde{\mathcal{R}}_m$). Since the points at which R_σ will be constructed will be in $[0, \hat{\alpha}]$ and the coefficients of R_σ converge to the respective coefficients of $\bar{R}$ as $\sigma \to 0$ we can also guarantee that $|R_\sigma(x)| \leq \frac{\mu_m(f)}{2}$ for $x \geq \hat{\alpha}$ when σ is sufficiently small. Thus, we need only work on $[0, \hat{\alpha}]$ and hence the proof given in [6] will apply. Finally, we observe that the constant error curve difficulty for varisolvent families as described in [2] does not occur here, since $\tilde{\mathcal{R}}_m$ is closed under scalar multiplication.

For the sufficiency, suppose $\bar{R}(x) = (\bar{p}_1 + \ldots + \bar{p}_m x^{m-1})/(\bar{q}x + 1)^m \in \tilde{\mathcal{R}}_m$ is nondegenerate, $\bar{q} > 0$ and $f(x) - \bar{R}(x)$ has $m + 2$ alternating extreme points at $\bar{y}_1, \ldots, \bar{y}_{m+2}$ where $0 \leq \bar{y}_1 < \bar{y}_2 < \ldots < \bar{y}_{m+2} \leq \alpha$ ($\bar{y}_{m+2}$ finite). If $\bar{R}$ is not a local best approximation, then for each $\delta_j > 0$ we can find $R_j(x) = (p_{1j} + \ldots + p_{mj}x^{m-1})/(q_j x + 1)^m \in \tilde{\mathcal{R}}_m$ such that $\|f - R_j\| < \|f - \bar{R}\|$, $|p_{ij} - \bar{p}_i| \leq \delta_j$, $i = 1, \ldots, m$ and $|q_j - \bar{q}| \leq \delta_j$. Let $\delta_j \to 0$ and let $\{R_j\}$ be a corresponding set of functions in $\tilde{\mathcal{R}}_m$ where we shall assume that $\|f - R_j\| \leq \|f - \bar{R}\|$ and $R_j \neq \bar{R}$ rather than $\|f - R_j\| < \|f - \bar{R}\|$. We shall show that this leads to a contradiction, proving our desired result and also that $\bar{R}$ is locally unique. For each j, let y_{ij} be a zero of $R_j - \bar{R}$ in $[\bar{y}_i, \bar{y}_{i+1}]$, $i = 1, \ldots, m + 1$. By going to subsequences, we may assume that $y_{ij} \to y_i^* \in [\bar{y}_i, \bar{y}_{i+1}]$ where we observe that $y_i^* = y_{i+1}^*$ is possible for some i, but $y_i^* = y_{i+1}^* = y_{i+2}^*$ can never occur. Similar equalities are possible for $\{y_{ij}\}_{i=1}^{m+1}$ for each j and if $y_{ij} = y_{i+1,j}$ for some i and j, then $R_j - \bar{R}$ and $R_j' - \bar{R}'$ vanish at y_{ij}. Suppose that for some i, $1 \leq i \leq m + 1$, $y_i^* = y_{i+1}^*$ ($= \bar{y}_{i+1}$). Then, if for some j $y_{ij} \neq y_{i+1,j}$ then by Rolle's Theorem there exists $y_{ij}' \in (y_{ij}, y_{i+1,j})$ such that $R_j'(y_{ij}') = \bar{R}'(y_{ij}')$. If $y_{ij} = y_{i+1,j} = \bar{y}_{i+1}$ then $R_j'(\bar{y}_{i+1}) = \bar{R}'(\bar{y}_{i+1})$ and we define $y_{ij}' = \bar{y}_{i+1}$ in this case. Thus, $y_{ij}' \to \bar{y}_{i+1}$ as $j \to \infty$. However, this implies that for sufficiently large j we must have $R_j \equiv \bar{R}$ by Lemma 2.4

which is our desired contradiction. Indeed, there are two possibilities to be considered.

CASE 1. $y_1^* < y_2^* < \ldots < y_{m+1}^*$. In this case we apply Lemma 2.4 with $m_1 = 0$ to $\bar{P}/\bar{Q}$ with respect to these points (i.e., $\bar{y}_i$ of Lemma 2.4 is y_i^*) with y_i of Lemma 2.4 set equal to y_{ij}, $i = 1, \ldots, m + 1$, j fixed, and $z_i = \bar{R}(y_{ij})$, $i = 1, \ldots, m + 1$. Then, for j sufficiently large we have that R_j satisfies the conclusion of Lemma 2.4 (i.e., coefficients of R_j sufficiently close to respective coefficients of $\bar{R}$ and $R_j(y_{ij}) = z_i$ with $|y_{ij} - y_i^*|$ and $|z_i - \bar{R}(y_i^*)|$ small). But $\bar{R}$ also satisfies the conclusion of Lemma 2.4 and since both $\bar{R}$ and $R_j \in \mathcal{R}_m$ we have by the uniqueness of Lemma 2.4 that $R_j \equiv \bar{R}$.

CASE 2. $y_{n_1}^* = y_{n_1+1}^*, \ldots, y_{n_\ell}^* = y_{n_\ell+1}^*$. In this case we apply Lemma 2.4 with $m_1 = \ell$ to the points $\{\tilde{y}_1, \ldots, \tilde{y}_{m-\ell+1}\}$ (a listing of the distinct points of $y_1^*, \ldots, y_{m+1}^*$) and the points $\{\bar{y}_{n_1}, \ldots, \bar{y}_{n_\ell}\}$. Letting u_i be the first index ν such that $y_{\nu k} \to \tilde{y}_i$ as $k \to \infty$, $i = 1, \ldots, m - \ell + 1$, we take the y_i of Lemma 2.4 as $y_{\mu_i k}$ (k fixed) and $z_i = \bar{R}(y_{\mu_i k})$, $i = 1, \ldots, m - \ell + 1$. We also take the y'_{ij} of Lemma 2.4 as $y'_{n_j k}$ (see definition just prior to case 1 of this proof) and $z'_j = \bar{R}'(y'_{n_j k})$, $j = 1, \ldots, \ell$. The desired result then follows immediately as in Case 1, with k playing the role that j did in Case 1. ∎

Corollary 2.7. Suppose $\bar{R}(x) = \bar{P}(x)/\bar{Q}(x) = (\bar{p}_1 + \ldots + \bar{p}_m x^{m-1})/(\bar{q}x + 1)^m \in \tilde{\mathcal{R}}_m$ is a local best approximation to $f(x)$ from $\tilde{\mathcal{R}}_m$ on $[0, \alpha]$ and $\bar{R}$ is nondegenerate with $\bar{q} > 0$. Then, $\bar{R}$ is locally unique.

Corollary 2.8. If $\bar{R}(x) = (\bar{p}_1 + \ldots \bar{p}_m x^{m-1})/(\bar{q}x + 1)^m \in \tilde{\mathcal{R}}_m$ is a best approximation to $f(x)$ from $\tilde{\mathcal{R}}_m$ on $[0,\alpha]$, $\bar{R}$ is nondegenerate and $\bar{q} > 0$, then $f(x) - \bar{R}(x)$ has at least $m + 2$ alternating extreme points.

The converse of Corollary 2.8 is most likely false since in the $m = 3$ case, $f(x) = e^{-x}$ and $\alpha = \infty$ we have essentially at least two best local approximations: $R_1(x) = (1.00805 - .27010x + .01447x^2)/(.27127x + 1)^3$ with error norm 8.05002×10^{-3}, achieved at the extreme points 0, .462, 2.178, 6.876 and 37.250 (with $e^{-x} - R_1(x) < 0$ at 0) and $R_2(x) = (.98663 + 2.52827x - .44972x^2)/(1.05109x + 1)^3$ with error norm 1.33720×10^{-2}, achieved at the extreme points 0, .172, .872, 2.950, 13.226 (with $e^{-x} - R_2(x) > 0$ at 0). These approximations were computed over a 20,001-point equally spaced grid imposed on [0, 40]. Although coefficients given above were rounded, using the actual coefficients computed the absolute errors at the extreme points in each case agreed to at least 15 significant figures. It seems very likely that a theoretical argument can be given starting with these two functions to show that at least two distinct local best approximations exist for this problem.

We can extend some of our results for $\mathcal{R}_m$ to the other possible configurations of the denominator of members of $\mathcal{R}_m$. If $m_1, \ldots, m_\ell > 0$ and $m_1 + \ldots + m_\ell = m$, we define $\tilde{\mathcal{R}}_{m_1, \ldots, m_\ell} = \{R = P/Q : P \in \Pi_{m-1}, Q(x) = (q_1 x + 1)^{m_1} \ldots (q_\ell x + 1)^{m_\ell}, 0 \leq q_1 < \ldots < q_\ell\}$. Although in general we cannot expect existence of best approximations from $\tilde{\mathcal{R}}_{m_1, \ldots, m_\ell}$ since the set of allowable coefficients is not closed, we have

Theorem 2.9. *Let $\ell > 0$, $m_1, \ldots, m_\ell > 0$, and $m_1 + \ldots m_\ell = m$. Then a nondegenerate $\bar{R}(x) = (\bar{p}_1 + \ldots + \bar{p}_m x^{m-1})/(\bar{q}_1 x + 1)^{m_1} \ldots (\bar{q}_\ell x + 1)^{m_\ell} \in \tilde{\mathcal{R}}_{m_1, \ldots, m_\ell}$ with $\bar{q}_1 > 0$ is a local best approximation to $f \in C[0, \alpha]$ on $[0, \alpha]$ from $\tilde{\mathcal{R}}_{m_1, \ldots, m_\ell}$ if and only if the error curve $E(x) = f(x) - \bar{R}(x)$ has at least $m+\ell+1$ alternating extreme points. (If $\alpha = \infty$, then we require $\lim_{x \to \infty} f(x) = 0$). Furthermore, in this case $\bar{R}$ is locally unique.*

The proof of this theorem requires only proving the analog of Lemma 2.4 for $\tilde{\mathcal{R}}_{m_1,\ldots,m_\ell}$. This proof is more involved than the proof of Lemma 2.4, but follows the same lines; the variable row of $H(t)$ turns out to be $((\bar{q}_1 t+1)\ldots(\bar{q}_\ell t+1),\ldots,t^{m-1}(\bar{q}_1 t+1)\ldots(\bar{q}_\ell t+1), -m_1 t(\bar{q}_2 t+1)\ldots(\bar{q}_\ell t+1)\bar{P}(t),$ $\ldots, -m_\ell t(\bar{q}_1 t+1)\ldots(\bar{q}_{\ell-1} t+1)\bar{P}(t))$.

As two consequences of this result we note first that if a nondegenerate best approximation $\bar{R}$ to f from $\mathcal{R}_m$ with all denominator coefficients positive is such that $f - \bar{R}$ has only $m + 2$ alternating extreme points, then $\bar{R} \in \tilde{\mathcal{R}}_m$; second, if a nondegenerate best approximation $\bar{R}$ to f from $\mathcal{R}_m$ has all its denominator coefficients positive and distinct, then $\bar{R}$ is actually the unique best approximation to f from $\mathcal{R}_m^{m-1}[0, \alpha] = \{R = P/Q : P \in \Pi_{m-1}, Q \in \Pi_m, Q > 0 \text{ on } [0, \alpha]\}$.

So far we have always constrained the numerator polynomial to lie in Π_{m-1}, but if we replace Π_{m-1} by Π_n and replace m by $n + 1$ in all expressions of numbers of alternating extreme points, then everything still goes through as long as $n < m$ or $\alpha < \infty$. If $n = m$ and $\alpha = \infty$ we conjecture that the results still go through if $\lim_{x\to\infty} f(x)$ exists and is finite; in this case "∞" may be an extreme point in the alternation theorems.

3. Results for $f(x) = e^{-x}$ and $\alpha = \infty$

In this section we describe the preceeding theory for the special case that $f(x) = e^{-x}$ and $\alpha = \infty$. It was this special case that motivated this general study and a report on this special case can be found in [4]. By the preceeding section we have that there exist best approximations to e^{-x} on $[0, \infty)$ from both $\mathcal{R}_m$ and $\tilde{\mathcal{R}}_m$. In addition, for the space $\tilde{\mathcal{R}}_m$ we have an alternation characterization of local best approximations and know that a local uniqueness result holds. As seen from the example at the end of the previous section, we conjecture that there may exist more

than one local best approximation in this case as well as at least one global best approximation. Whether or not there is precisely one global best approximation is not known. Finally, we conjecture that there is a best approximation to e^{-x} from $\mathcal{R}_m$ which is actually in $\tilde{\mathcal{R}}_m$. In fact, we believe that each best approximation to e^{-x} from $\mathcal{R}_m$ is in $\tilde{\mathcal{R}}_m$ (if more than one exists). We have proved this stronger statement in the case that $m = 2$ [4]. Also, observe that the numerical results given in [4] support this conjecture.

4. Computations

Our algorithms for computing approximations from $\mathcal{R}_m$ and $\tilde{\mathcal{R}}_m$ involve linearizing the denominator by Taylor's theorem and setting up an iterative procedure, using a combination Remes-differential correction algorithm to compute an approximation at each inner stage. For $\mathcal{R}_m$ set $g(q_1, \ldots, q_m, x) = \prod_{i=1}^{m} (q_i x + 1)$ and define $\psi_j(q_1, \ldots, q_m, x) = \sum_{\substack{i=1 \\ i \neq j}}^{m} \prod (q_i x + 1)$ for $j = 1, \ldots, m$, $\psi_0(q_1, \ldots, q_m, x) = g(q_1, \ldots, q_m, x) - \sum_{\nu=1}^{m} q_\nu \psi_\nu (q_1, \ldots, q_m, x)$. If $\bar{R}(x) = \bar{P}(x) / \prod_{i=1}^{m} (\bar{q}_i x + 1)$, $0 \leq \bar{q}_1 \leq \bar{q}_2 \leq \ldots \leq \bar{q}_m$ is an approximation to $f(x)$ at some step in the algorithm, then a new approximation $R(x) = (p_0 + p_1 x + \ldots + p_m x^{m-1}) / \prod_{i=1}^{m} (q_i x + 1)$ is found by calculating $p_0, \ldots, p_{m-1}, q_1, \ldots, q_m$ to minimize

$$\| f(x) - (p_0 + p_1 x + \ldots + p_m x^{m-1}) / (q_1 \psi_1(\bar{q}_1, \ldots, \bar{q}_m, x) + \ldots + q_m \psi_m(\bar{q}_1, \ldots, \bar{q}_m, x) + \psi_0(\bar{q}_1, \ldots, \bar{q}_m, x)) \|$$

over a finite subset T of $[0, \alpha]$, with the restrictions $0 \leq \bar{q}_1 \leq \bar{q}_2 \leq \ldots \leq \bar{q}_m \leq \beta$ (where β depends on the approximation desired). The ordering restrictions $\bar{q}_1 \leq \bar{q}_2 \leq \ldots \leq \bar{q}_m$ were found to be necessary to obtain convergence. Observe that the denominator in this problem is

precisely the linearization of $g(q_1, \ldots, q_m, x)$ via Taylor's theorem applied to the first m independent variables. The $\tilde{\mathcal{R}}_m$ algorithm uses the same approach; the linearized denominator for this algorithm is $qmx(\bar{q}x + 1)^m + [(1 - m)\bar{q}x + 1](\bar{q}x + 1)^{m-1} \equiv q\psi_1(\bar{q}, x) + \psi_0(\bar{q}, x)$.

If α is a large finite number or $\alpha = \infty$ (in which we case we consider $[0, \hat{\alpha}]$ instead of $[0, \infty)$ for some large finite $\hat{\alpha}$), and we wish to use a fairly fine mesh in order to get an accurate approximation over $[0, \alpha]$, then card(T) will be large. Since this leads to a large and difficult linear programming problem and can cause storage problems in the differential correction algorithm we used the Remes-Difcor algorithm [3] for calculating the linearized minimum. This algorithm applies the differential correction algorithm to certain small subsets of T chosen in such a manner (depending on alternation) that convergence to the solution on T occurs. Thus, we had no a priori guarantee that this would work since a standard alternation theory has not been developed for this problem; however, in most cases both inner and outer algorithms converged and we obtained approximations satisfying theorem 2.9. Although a precise study of these algorithms remains to be done, we conjecture that the $\tilde{\mathcal{R}}_m$ algorithm will converge (assuming the convergence of the inner iterations) if the initial guess for the denominator coefficient is sufficiently good. We make no such conjecture for the $\mathcal{R}_m$ algorithm as presently constituted, since if $\bar{q}_i = \bar{q}_{i+1}$ at some stage, then $\psi_i(\bar{q}_1, \ldots, \bar{q}_m, x) \equiv \psi_{i+1}(\bar{q}_1, \ldots, \bar{q}_m, x)$, and the q_i and q_{i+1} at the next stage will not be uniquely determined. In practice the Remes-Difcor algorithm has chosen q_i and q_{i+1} at the next stage so that either $q_i = q_{i-1}$ (or $q_i = 0$ if $i = 1$) or $q_{i+1} = q_{i+2}$ (or $q_{i+1} = \delta$ if $i = m - 1$).

As an example, consider the problem of approximating the function f on $[0, 20]$ by functions of the form $\dfrac{p_1 + p_2 x}{(q_1 x + 1)(q_2 x + 1)(q_3 x + 1)}$ with

$0 \le q_1 \le q_2 \le q_3 \le 10$, where f is defined by

$$f(x) = \begin{cases} 6.7x - 9, & 0 \le x \le 1 \\ \frac{29}{30}x - \frac{49}{15}, & 1 \le x \le 2 \\ -\frac{4}{3}, & 2 \le x \le 4 \\ \frac{173}{300}x - 3.64, & 4 \le x \le 8 \\ -\frac{661}{2700}x + \frac{1979}{675}, & 8 \le x \le 16 \\ \frac{32,653}{65,340}x - \frac{29,341}{32,670}, & 16 \le x \le 20 . \end{cases}$$

Applying the $\mathcal{R}_m$ algorithm with m = 3, replacing Π_2 in the numerator by Π_1, we obtained $R_1(x) = \frac{1.00000x - 10.00000}{(.25000x+1)(.50000x+1)^2}$. In accordance with Theorem 2.9 and the remarks following that theorem there were five alternating extreme points; these occurred at 0, 4, 8, 16 and 20, with $f(0) - R_1(0) = 1.000000000$. This was obtained using the actual denominator coefficients as the initial guess; using instead the initial guess $\bar{q}_1 = 0.33$, $\bar{q}_2 = \bar{q}_3 = 0.55$ produced $q_1 = .24637$, $q_2 = q_3 = .50256$ after one iteration, but the next iteration produced $q_1 = q_2 = .25000$, $q_3 = .75000$, the following iteration produced $q_1 = q_2 = .31250$, $q_3 = .62500$, and the algorithm failed to converge after 12 iterations. Starting with initial guess $\bar{q}_1 = \bar{q}_2 = .31969$, $\bar{q}_3 = .50216$ (these were obtained by running the algorithm for 12 iterations with initial guess $\bar{q}_1 = 0.1$, $\bar{q}_2 = 0.4$, $\bar{q}_3 = 0.7$), after 4 iterations we obtained the local best approximation $R_2(x) = \frac{1.00011x - 10.00000}{(.31578x + 1)^2(.62556x + 1)}$; the extreme points were 0, 4, 8, 16 and 20 with $f(0) - R_2(0) = 1.000002717$.

We also approximated the same f on [0, 20] using the $\tilde{\mathcal{R}}_m$ algorithm with m = 3, again replacing Π_2 in the numerator by Π_1. Using either initial guess $\bar{q} = 0.2$ (this required 7 iterations) or $\bar{q} = 0.6$ (6 iterations) we obtained the local best approximation $R_3(x) = \frac{.92290x - 9.22429}{(.38772x + 1)^3}$; the alternating extreme points were 4, 8, 16 and 20, with $f(4) - R_3(4) = -1.000011692$.

Using as initial guess $\bar{q} = 1.4$ (8 iterations), $\bar{q} = 1.8$ (11 iterations), $\bar{q} = 4.0$ (8 iterations), $\bar{q} = 6.0$ (8 iterations) or $\bar{q} = 8.0$ (9 iterations) we obtained the local best approximation $R_4(x) = \frac{-112.16689x - 7.94741}{(2.69005x + 1)^3}$; the alternating extreme points were 0, .12, 4 and 8, with $f(0) - R_4(0) = -1.052593329$.

All computations were done on a UNIVAC 1106 (which has roughly 18 decimal digits of accuracy in double precision), and for each of the functions R_1, R_2, R_3 and R_4 the absolute values of the error at the extreme points agreed at least to the accuracy printed out (10 significant figures). For further numerical results see [4].

5. Summary

For convenience we list some open questions in this last section, some of which were mentioned earlier.

1. Conjecture: The best approximation to e^{-x} on $[0, \alpha]$ from $\mathcal{R}_m$ is actually in $\tilde{\mathcal{R}}_m$ (for all $\alpha > 0$ or for all α sufficiently large and $\alpha = \infty$).

2. Characterize those functions for which the best approximation from $\mathcal{R}_m$ is actually in $\tilde{\mathcal{R}}_m$.

3. Compute the constant of geometric convergence for $\operatorname{dist}(e^{-x}, \mathcal{R}_m)$ on $[0, \infty)$; that is, find $q > 1$ such that $\overline{\lim_{m\to\infty}} [\operatorname{dist}(e^{-x}, \mathcal{R}_m)]^{1/m} = \frac{1}{q}$. Is $q = \tilde{q}$ where $\tilde{q}$ is the geometric constant defined by $\overline{\lim_{m\to\infty}} [\operatorname{dist}(e^{-x}, \frac{P(x)}{(1 + \frac{x}{m})^m})]^{1/m} = \frac{1}{\tilde{q}}$ where $P(x)$ ranges over Π_{m-1} and $0 \leq x < \infty$ (see [7]).

4. What is the situation with regard to alternation if a local best approximation is degenerate? For example, suppose $\bar{R}(x) = \dfrac{\bar{p}_1 + \bar{p}_2 x}{(\bar{q}x + 1)^2}$ with $\bar{q} > 0$ and $\bar{q}x + 1$ is not a factor of $\bar{p}_1 + \bar{p}_2 x$. Then 4 alternating extreme points are necessary if $\bar{R}$ is to be a local best approximation to f from $\tilde{\mathcal{R}}_3$, but probably not sufficient. Are 5 alternating extreme points necessary and/or sufficient?

5. What is the situation when $n \geq m$ and $\alpha = \infty$ (see the conjecture at the end of section 2)?

6. Conjecture: If the denominator coefficient is chosen sufficiently close to that of a local best approximation from $\tilde{\mathcal{R}}_m$, then the $\tilde{\mathcal{R}}_m$ algorithm will converge to it (assuming the inner iterations converge).

7. How many local best approximations are there?

8. When will global uniqueness occur?

Finally, we would like to thank Professor R. S. Varga for bringing [5] to our attention, where some of the results of this paper and [4] were proved independently.

REFERENCES

1. Cody, W. J., G. Meinardus and R. S. Varga. Chebyshev rational approximations to e^{-x} on $[0, \infty)$ and applications to heat-conduction problems, J. Approx. Theory, 2 (1969), 50-65.

2. Dunham, C. B. Necessity of alternation, Canad. Math. Bull., 10 (1967), 743-744.

3. Kaufman, E. H., Jr., D. J. Leeming and G. D. Taylor, A combined Remes-Differential correction algorithm for rational approximation, to appear Math. Comp.

4. Kaufman, E. H., Jr. and G. D. Taylor. Best rational approximations with negative poles to e^{-x} on $[0, \infty)$, Padé and Rational Approximations: Theory and Applications, Academic Press, Inc., New York, 1977.

5. Lau, Terence, C-Y, Rational exponential approximation with real poles, preprint.

6. Rice, J. R. Approximation of Functions, Vol. II, Addison-Wesley, Reading, Mass., 1969.

7. Saff, E. B., A. Schönhage and R. S. Varga. Geometric convergence to e^{-z} by rational functions with real poles, Numer. Math, 25 (1976), 307-322.

8. Saff, E. B. and R. S. Varga. Angular overconvergence for rational functions converging geometrically on $[0, \infty)$, Theory of Approximations with Applications, pp. 238-256, Academic Press, Inc., New York, 1976.

E. H. Kaufman, Jr.
Department of Mathematics
Central Michigan University
Mount Pleasant, Michigan 48859

G. D. Taylor
Department of Mathematics
Colorado State University
Fort Collins, Colorado 80523